EDGEWOOD

CHEMICAL BIOLOGICAL CENTER

U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND

ECBC-TR-444

RESOLUTION OF THE SPATIAL DISTRIBUTION OF THE BACTERIOPHAGE MS2 PROTEIN AND RNA COMPONENTS

Deborah A. Kuzmanovic

GEO-CENTERS

GEO-CENTERS, INC.-GUNPOWDER BRANCH

Ilya Elashvili
Charles H. Wick

RESEARCH AND TECHNOLOGY DIRECTORATE

Catherine O'Connell

NIST BIOTECHNOLOGY DIVISION
Gaithersburg, MD 20899-8311

Susan Krueger

NIST CENTER FOR NEUTRON RESEARCH
Gaithersburg, MD 20899-8562

May 2005

ABERDEEN PROVING GROUND, MD 21010-5424

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
XX-05-2006	Final	May 2001 - Dec 2003

4. TITLE AND SUBTITLE
Resolution of the Spatial Distribution of the Bacteriophage MS2 Protein and RNA Components

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kuzmanovic, Deborah A. (GEO-CENTERS, INC.); Elashvili, Ilya; Wick, Charles H. (ECBC); O'Connell, Catherine (NIST-Biotechnology); and Krueger, Susan (NIST-CNR)

5d. PROJECT NUMBER
None

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
DIR, ECBC, ATTN: AMSRD-ECB-RT-DD, APG, MD 21010-5424 GEO-CENTERS, INC., Gunpowder Branch, P.O. Box 68, APG, MD 21010-0068
NIST Biotechnology Division, 100 Bureau Drive, Stop 8311, Gaithersburg, MD 20899-8311
NIST Center for Neutron Research, 100 Bureau Drive, Stop 8562, Gaithersburg, MD 20899-8311

8. PERFORMING ORGANIZATION REPORT NUMBER
ECBC-TR-444

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Small-angle neutron scattering (SANS) and the contrast variation technique were used to determine the spatial relationship of the individual components of the MS2 virion (protein shell and genomic RNA). Our results show that (a) the MS2 RNA is tightly compacted within the virion, confined to a radius of 83 ± 1Å, and (b) the MS2 coat protein shell extends to a radius of 136 ± 2Å, and has a thickness of 21 ± 1Å. Therefore, the protein core is composed primarily of solvent, and the RNA core is compact.

15. SUBJECT TERMS

Small Angle Neutron Scattering (SANS)	Virus structure	Virus	Detection
Purification	Virus detection	MS2	Bacteriophage

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Sandra J. Johnson
U	U	U	UL	21	19b. TELEPHONE NUMBER *(include area code)* (410) 436-2914

Blank

PREFACE

This work was started in May 2001 and completed in December 2003.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for public release. Registered users should request additional copies from the Defense Technical Information Center; unregistered users should direct such requests to the National Technical Information Service.

Blank

CONTENTS

FIGURES

TABLES

RESOLUTION OF THE SPATIAL DISTRIBUTION OF THE BACTERIOPHAGE MS2 PROTEIN AND RNA COMPONENTS

1. INTRODUCTION

Bacteriophage MS2 is a 275Å RNA virus that infects male *Escherichia coli* (Stockley *et al.*, 1994). Because of its small size, relatively simple composition, and ease of growth, MS2 is used as a model organism for a number of macromolecular (Peabody and Al-Bitar, 2001; Stockley *et al.*, 1994). A great deal is known about the MS2 bacteriophage. Its complete genome has been sequenced (Fiers *et al.*, 1976). The 3,569 nucleotide genome encodes a coat protein, a maturation protein (or A protein), a replicase subunit, and a lysis protein (Atkins *et al.*, 1979; Fiers *et al.*, 1976). The MS2 coat protein is the primary structural component of the MS2 protein shell. In addition to this function, it binds to the MS2 operator site and acts as a translation repressor of transcription of the MS2 replicase cistron. The A protein (or maturation protein) has been shown to be involved in attachment to the bacterial pilus, replication, RNA packing and infectivity *in vivo*. The replicase and lysis proteins are involved in replication and the lysis of the E. coli bacteria, respectively.

The MS2 virion is comprised of three components: the coat protein (relative molecular weight Mr = 13,700), the A protein (relative molecular weight Mr = 44,000), and a single-stranded RNA molecule. The 3-dimensional structure of the intact virion has been determined and refined at 2.8Å resolution (Golmohammadi *et al.*, 1993; Valegard *et al.*, 1991; Valegard *et al.*, 1990; Valegard *et al.*, 1986). From crystallographic analysis, the MS2 virion is thought to be composed of 90 coat protein homo-dimers arranged in a quasi-equivalent T = 3 lattice to form the icosahedral capsid shell of the type described by Caspar and Klug, 1962. In the capsid, coat protein dimers are thought to adopt two possible non-covalent quasi-equivalent arrangements, A/B and C/C. The A and C subunits interact at the quasi 6-fold axes, while the B-type subunits interact at the 5-fold axes. Structurally, the primary difference between these conformers lies in the position of the FG loop region of the protein. In the A and C subunits, the FG loop is extended, while in the C subunit it is folded back in the direction of the protein (Valegard *et al.*, 1991; Valegard *et al.*, 1990; Valegard *et al.*, 1986). In addition to the coat protein dimers, the MS2 capsid contains a single copy of the A protein. The A protein also has been shown to be tightly associated with the MS2 genomic RNA (Shiba and Suzuki, 1981), which is important for RNA packing *in vitro* (Argetsinger, 1966 #63; Heisenberg, 1966 #64). In addition to its role in RNA packing, the A protein is important for host recognition, attachment, and subsequent transfer of phage genomic RNA into its host (Stockley *et al.*, 1994).

The purpose of this study is to extend the structural characterization of the MS2 phage by examining its physical characteristics in solution by use of small angle neutron scattering (SANS). Specifically, we are interested in determining the spatial relationship between the genomic RNA to its protein shell.

The use of SANS for determining molecular weight has been described in detail (Jacrot and Zaccai, 1981). In general, SANS is a process where a neutron beam is passed through a sample, and the resulting scattering pattern reveals information about the average size,

shape, and orientation of the sample (Svergun and Koch, 2002; Glinka *et al.*, 1998). SANS experiments do not cause radiation damage to the sample, and typical experiments can be performed under physiological conditions in solution. Also, when the concentration (or particle number) of the sample is known, then the molecular weight of the sample can be determined by SANS since the data are obtained on an absolute scale (usually in cm^{-1}). Similarly, if the total molecular weight of the sample is known, then the concentration of the particles in the sample can be determined (Mazzone, 1998). A number of phage and viral molecular weights have been successfully determined by this method (e.g., Frog virus 3 Cuillel *et al.*, 1979), influenza (Cusack *et al.*, 1985), pf1 phage (Torbet, 1979), and Semiliki Forest virus (Freeman and Leonard, 1980).

SANS is a powerful tool for structural analysis but, when combined with the contrast variation method, it also permits additional structural information to be obtained about the individual components in a macromolecular complex. In the case of MS2, the contrast variation technique involves varying the solvent water to deuterated water ratio so that structural information about the protein and nucleic acid components can be obtained separately (Glinka *et al.*, 1998; Struhrmann and Miller, 1978).

2. EXPERIMENTAL PROCEDURES*

2.1 Bacteriophage, Hosts, and Medium.

The MS2 bacteriophage strain 15597-B1 and its *Escherichia coli (E. coli)* host 15597 were purchased from the American Type Culture Center (Manassas, VA). *E. coli* strain 15597 was grown on MS2 broth. MS2 broth contains, per liter: 10 g tryptone, 8 g NaCl, and 1 g Bacto-yeast. After autoclaving: 10 mL of sterile 10% glucose, 2 mL of 1 mol/l (M) $CaCl_2$, and 10 mg/mL of thiamine hydrochloride were added per liter (Davis and Sinsheimer, 1963). The MS2 was stored in Tris-Salt-Magnesium (TSM) buffer unless otherwise stated. The TSM buffer contains 10 mM Tris (pH 7.0), 100 mM NaCl, and 1 mM $MgCl_2$.

2.2 Growth of Bacteriophage.

The MS2 phage was grown using protocols modified from (Sambrook and Russell, 2001) and is described below. Fresh MS2 broth (3 mL) was inoculated with 400 µl of an overnight culture of *Escherichia coli* strain 15597 [American Type Culture Center (Manassas, VA)] at $OD_{600} = 1$ (1×10^9 cells/mL). The MS2 bacteriophage was added to the inoculant at a multiplicity of infection of 0.01 to 3 and incubated for 20 min at 37 °C. The mixture was added to 500 mL of pre-warmed MS2 broth and incubated 8-12 hr. Cell lysis was induced by adding 20 mL of chloroform followed by shaking for 10 min at 37 °C. Cultures were cooled to room

*Certain commercial materials, instruments, and equipment are identified in this manuscript to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology (NIST) nor does it imply that the materials, instruments, or equipment identified is necessarily the best available for the purpose.

temperature and then DNAse I and RNAse were each added to a final concentration of 50 mg/mL. The cultures were incubated for 30 min at room temperature, and then 29.2 g of NaCl (final concentration, 1M) was added. The mixture was incubated for 1 hr on ice and centrifuged at 107,910 m/s^2 (11,000 x g) for 10 min at 4 °C. To the supernatant, 45.2 g of ammonium sulfate was added to produce a 20% (w/w) saturated solution and incubated at 4 °C for 2 hr. The mixture was centrifuged at 11,000 x g for 30 min at 4 °C. Ammonium sulfate (75.2 g) was added to the supernatant to produce a 50% (w/w) saturated solution. Following overnight incubation at 4 °C, the sample was centrifuged at 11,000 x g for 20 min at 4 °C. The pellet was resuspended in 30 mL TSM, and 11.0 g of ammonium sulfate was added. The solution was incubated overnight at 4 °C and then centrifuged at 11,000 x g for 20 min at 4 °C. The pellet was resuspended in 35 mL TSM and centrifuged at 11,000 x g for 30 min. For the final precipitation step, the bacteriophage mixture was incubated at 4 °C for 1 hr or overnight and centrifuged at 11,000 x g for 10 min at 4 °C. The aqueous phase contained the crude phage particles.

Purified MS2 phage was isolated by cesium chloride equilibrium gradient. The cesium chloride protocol used was as described by Sambrook and Russell, 2001 with the following modifications. Cesium chloride was dissolved in TSM medium. Ultra-centrifugation was performed at 23 °C. Following centrifugation, the samples were transferred to Slide-a Lyzer (Pierce, Rockford, IL) and dialyzed in 500 mL TSM for 24 hr with 2 changes of buffer. The measured density of the MS2 particles was 1.38 + 0.01 g/cm^3, which is the same density value reported by Strauss and Sinsheimer, 1963. Samples for SANS measurements were made in TSM buffers containing 0, 10, 65, 85, and 100% D_20. The samples were dialyzed in the appropriate buffers for 2 hr at room temperature, with two changes of buffer, and then transferred to sample holders.

2.3 SANS Measurements.

SANS measurements were performed on the 30-m SANS instruments at the NIST Center for Neutron Research in Gaithersburg, MD (Glinka *et al.*, 1998). The neutron wavelength, λ, was 6Å, with a wavelength spread, $\Delta\lambda/\lambda$, of 0.15. Scattered neutrons were detected with a 64 cm x 64 cm 2-dimensional position sensitive detector with 128 x 128 pixels. Raw counts were normalized to a common monitor count and corrected for empty cell counts, ambient room background counts and non-uniform detector response. Data were placed on an absolute scale by normalizing the scattered intensity to the incident beam flux. Finally, the data were radially averaged to produce scattering intensity, I(Q), versus Q curves, where $Q = 4\pi\sin(\theta)/\lambda$, and 2θ is the scattering angle. Sample-to-detector distances of 12 m and 2.5 m were used to cover the range $0.005\ Å^{-1} \leq Q \leq 0.17Å^{-1}$. The scattered intensities from the samples were then further corrected for buffer scattering and incoherent scattering from hydrogen in the samples.

2.4 SANS Data Analysis.

The Guinier approximation, $I(Q) = I(0)\exp(-Q^2Rg^2/3)$, was used on the low-Q portions of the data to obtain initial values for the radius of gyration, Rg, and the forward scattering intensity, I(0), of the samples. This analysis is valid only in the region where $QRg \sim 1$. The GNOM program (Semenyuk and Svergun, 1991), which uses all of the data, rather than a limited data set at small Q values, was used to determine the distance distribution function, P(r); the radius of gyration, Rg; the forward scattering intensity, I(0); and the maximum dimension, D_{max}. Since all of the data are used, this approach typically leads to more accurate determinations of Rg and I(0) that are less influenced by possible aggregation effects.

Since MS2 can be approximated very well by a spherical shell at the resolution level of the SANS measurements, the data were also fit to a core-shell sphere model (Semenyuk and Svergun, 1991) to obtain the radius of the protein shell and RNA core. The neutron scattering length density of the RNA core was an additional fitting parameter that allowed the amount of water, versus RNA, in the core to be calculated using eq. 1:

$$\rho_{CORE} = X\rho_{RNA} + (1-X)\,\rho_{SOLVENT} \qquad (1)$$

where X is the fraction of RNA in the core, ρ_{CORE} is the fitted scattering length density of the core portion of the core-shell model, and ρ_{RNA} and $\rho_{SOLVENT}$ are the known scattering length densities of the RNA and the solvent, respectively. The core-shell model fits consider the resolution function of the SANS instruments.

The scattered intensities from the MS2 protein/RNA complex were decomposed into the scattering from their components, $I_{PROT}(Q)$ and $I_{RNA}(Q)$ using eq. 2:

$$I(Q) = \Delta\rho_{PROT}^2 I_{PROT}(Q) + \Delta\rho_{PROT}\Delta\rho_{RNA} I_{PROTRNA}(Q) + \Delta\rho_{RNA}^2 I_{RNA}(Q) \qquad (2)$$

where $\Delta\rho = (\rho-\rho_s)$ is the contrast, or the difference between the scattering length density of the molecule (ρ) and the solvent (ρ_s). The cross-term, $I_{PROTRNA}(Q)$, represents the interference function between the protein and RNA components. The known quantities in eq. 1 are $\Delta\rho_{PROT}$ and $\Delta\rho_{RNA}$. The unknowns are $I_{PROT}(Q)$, $I_{RNA}(Q)$, and $I_{PROTRNA}(Q)$. Since measurements were made at five different contrasts, or D_2O/H_2O buffer conditions, there is sufficient information to solve for the three unknown component intensities from the set of simultaneous equations for I(Q) at each contrast.

The Mw values of the protein and RNA components of MS2 were calculated in a similar manner using eq. 3:

$$I(0) = n(\Delta\rho_{PROT}V_{PROT} + \Delta\rho_{RNA}V_{RNA})^2 \qquad (3)$$

where n is the number density of MS2 particles, and V_{PROT} and V_{RNA} are the volumes of the protein and RNA components, respectively. These volumes can be written as $V = M_w/(N_A d)$, where d is the mass density and N_A is Avogadro's number. Now, eq. 2 can be rewritten as:

$$\left[\frac{I(0)}{n}\right]^{1/2} = \left(\frac{|\Delta\rho_{PROT}|}{N_A\, d_{PROT}}\right) M_{WPROT} + \left(\frac{|\Delta\rho_{RNA}|}{N_A\, d_{RNA}}\right) M_{WRNA} \qquad (4)$$

where $d_{PROT} = 1.38$ g/cm^3 and $d_{RNA} = 1.89$ g/cm^3. Now, there are only two unknowns, $\Delta\rho_{PROT}$ and $\Delta\rho_{RNA}$. The I(0) values obtained from the GNOM analysis of the data for each D_2O/H_2O buffer can be used with the measured number densities to solve the set of simultaneous equations for these two unknowns to obtain the M_w values for the protein and RNA components separately in the MS2 complex. The total M_w value is then simply the sum of the two component M_w values. It is important to note that I(0) must be on an absolute scale, usually in cm^{-1}, to obtain accurate M_w values from eq. 3 or 4.

2.5 Number Density Determinations.

Number density determinations were made using two methods: (1) the concentration was measured by optical density (OD) using a conventional spectrophotometer and then the number density was estimated using this information, and (2) the number density was obtained directly using the Integrated Virus Detection System (IVDS), which is a particle counting method. Measurements were obtained both before and after dialysis and subsequent SANS experiments. However, only the measurements taken after dialysis are used and reported here.

The number densities calculated from the OD_{260} measurements were found from the measured concentration, c, using $n = cN_A/Mw$, where N_A is Avogadro's number, and Mw is the total molecular weight of the MS2 particle. Since n has units of cm^{-3}, c must be converted to units of g·cm^{-3}. Sample concentrations were measured after the SANS experiments by measuring the absorbance at 260 nm and using Beer's Law,

$$c = A_{260}/\varepsilon \cdot L, \qquad (5)$$

where ε is the molar coefficient, and L is the pathlength of the light, to calculate the concentration (Eisenberg, 1979). Since the molar coefficient is also dependent upon the total Mw of the particle, this method of determining the number density is only useful if the total Mw of the particle is known (Eisenberg, 1979). Sample concentrations were measured in duplicate using a Hewlett-Packard model 8450A spectrophotometer. The spectrophotometer was calibrated using NIST transmittance and wavelength Standard Reference Material numbers 930, 2031, and 2034.

The IVDS was used to determine MS2 particle number directly. The IVDS is a bipartite instrument consisting of (1) an ultra-filtration unit for use in the purification and/or concentration of materials for analysis, and (2) a gas-phase electrophoretic mobility analyzer (GEMMA) detector for particle counting and sizing measurements (Wick and McCubbin, 1999a). The ultra-filtration unit has been previously described (Wick and McCubbin, 1999c) and was not used for these experiments. This work used the GEMMA detector system only. The GEMMA detector consists of an electro spray that sprays the sample into the detector, a differential mobility analyzer to separate the sample by size and a condensation particle counter

for particle counting. These components are in a single module. The complete IVDS instrument has been previously described in detail and was originally designed to detect, quantify and size viruses in the 10 to 100 nm size range (Wick and McCubbin, 1999a; Wick and McCubbin, 1999b; Wick and McCubbin, 1999c). The IVDS instrument was calibrated using a NIST-Traceable Standard Reference Material. The particle number of samples for the experiments described here was determined using the optimal usage procedures and calibration conditions described by Kuzmanovic (unpublished data, 2003).

Table 1. Parameters from Core-Shell Model Fit

Sample	R1 (Å)	R2 (Å)	t (Å)	I(0) cm^{-1}
		Experiment #1		
0% D_2O	112	134	22	4.03 ± 0.02
10% D_2O	84	141	57	2.50 ± 0.02
65% D_2O	118	137	19	0.46 ± 0.01
85% D_2O	115	133	18	2.10 ± 0.01
100% D_2O	115	139	24	6.65 ± 0.05
Average Values*	115 ± 1	136 ± 1	21 ± 1	—
		Experiment #2		
0% D_2O	110	136	26	5.47 ± 0.01
10% D_2O	95	139	44	4.46 ± 0.01
65% D_2O	116	144	28	0.79 ± 0.03
85% D_2O	112	136	14	2.83 ± 0.02
100% D_2O	112	139	27	7.94 ± 0.02
Average Values*	113 ± 1	139 ± 2	24 ± 3	—

*The values obtained for the 10% D_2O samples were excluded when calculating the average values. Errors in the average values for R1, R2, and t are standard deviations of the mean and do not include unknown systematic errors. Errors in I(0) values are standard deviations of the mean from several fits that were made for each %D_2O condition, yielding the same values for R1, R2, and t.

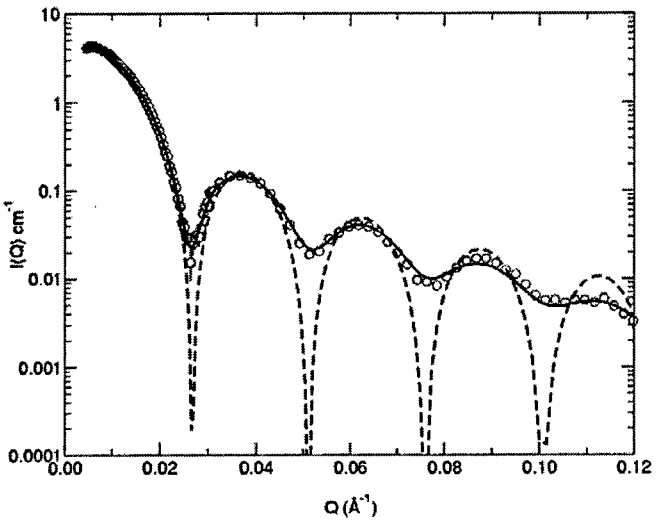

Figure 1. Core-Shell Model Fit for MS2. A sample core-shell model fit, with (—) and without (– – –) correcting the model for instrumental resolution effects, for the 100% D_2O data (○) of Experiment #1.

3. RESULTS

Two complete contrast variation series of measurements were performed on two different MS2 sample preparations and are designated as Experiment #1 and Experiment #2. At the resolution of the SANS measurements, the shape of an MS2 particle can be approximated very well by a spherical shell, with inner radius, R1, outer radius, R2, and shell thickness, t = R2 – R1. Fits to this core-shell model were made for the data at each contrast, and the results are shown in Table 1. A sample model fit, made with and without correcting the model for instrumental resolution effects, is shown for the 100% D_2O data of Experiment #1 in Figure 1. Table 1 shows that, in all cases, the outer radius of the shell, R2, consistently falls between the values of 134Å and 144Å. The (core) inner radius, R1, falls between 110Å and 118Å, except for the 10% D_2O buffer sample, which consistently shows a much smaller R1 value for both experiments. The RNA in the core scatters strongly in comparison with the protein shell under these solvent conditions. Thus, at this contrast, the lower value for R1 could be an indication that the RNA is actually packed compactly and does not completely fill the core region, with the remainder of the core being mostly solvent. The amount of water in the core was calculated from the fitted scattering length density of the core region for the data at each contrast using eq. 1. The average fraction of water in the core region was 0.81 ± 0.04 for Experiment #1 and 0.77 ± 0.03 for Experiment #2. If the values obtained in 10% D_2O are excluded, the average parameters obtained from the core-shell model fit are R1 = 115 ± 1, R2 = 136 ± 1, and t = 21 ± 1 for Experiment #1, and R1 = 113 ± 1, R2 = 139 ± 2, and t = 24 ± 3 for Experiment #2.

Table 2. Parameters from Distance Distribution Function Determination

Sample	Rg (Å)	I(0) cm^{-1}
	Experiment #1	
0% D_2O	114 ± 1	4.10 ± 0.02
10% D_2O	114 ± 1	2.60 ± 0.02
65% D_2O	130 ± 1	0.46 ± 0.07
85% D_2O	120.0 ± 0.05	2.10 ± 0.01
100% D_2O	122.0 ± 0.05	6.70 ± 0.01
	Experiment #2	
0% D_2O	115 ± 1	5.48 ± 0.03
10% D_2O	114 ± 1	4.48 ± 0.03
65% D_2O	127 ± 1	0.77 ± 0.01
85% D_2O	124.0 ± 0.05	2.82 ± 0.02
100% D_2O	123.0 ± 0.05	7.94 ± 0.02

Errors are standard deviations of the mean for several equally good fits made at each %D_2O condition.

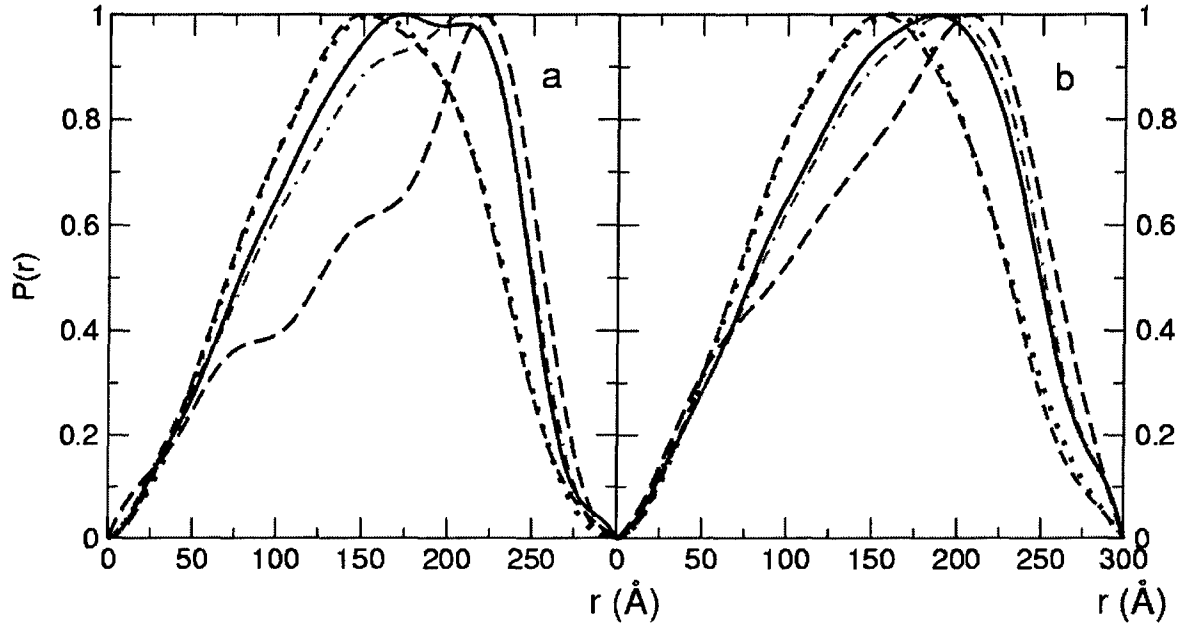

Figure 2. Distance Distribution Functions for MS2. Distance distribution functions, P(r) vs. r, from the data for samples in 100% D_2O (—), 85% D_2O (· – –), 65% D_2O (– –), 10% D_2O (– – –), and 0% D_2O (· · ·) from (a) Experiment #1 and (b) Experiment #2.

Distance distribution functions, P(r), were obtained from the data as described in Section 2, and the fitted parameters are listed in Table 2. The resultant P(r) functions for Experiment #1 are plotted in Figure 2a, and those for Experiment #2 are plotted in Figure 2b. The P(r) functions are normalized so that the peak value is equal to 1.0 in each case. The maximum distance, D_{max}, in all cases, was 300Å, which is larger than 2*R2. By definition, D_{max} is the distance at which P(r) goes to zero. Thus, D_{max} suggests a sharp boundary between the particle and its surroundings. Since the shape of the MS2 coat protein region is actually icosahedral, this boundary is not sharp, and the P(r) functions suggest that the particle does actually extend beyond 2*R2. However, the number of probable distances beyond R2 drops sharply.

For both experiments, the P(r) function for the 65% D_2O sample is consistent with that of a hollow spherical shell. In this case, the peak of the distance distribution is at ~200Å, consistent with the fact that the most probable distances are occurring beyond 2*R1. In 65% D_2O, the scattering length density of the RNA component of the complex is the same as that of the solvent. Thus, the scattering from the RNA is masked and only the scattering from the protein is observed. Therefore, the fact that MS2 most closely resembles a hollow sphere at this contrast is expected.

On the other hand, the peak of the distance distribution function occurs at values smaller than 2*R1 for the data obtained at the other contrasts. The scattering from the RNA component is the strongest, relative to the protein component, in the 0% and 10% D_2O conditions, which is evident from the fact that the peak in P(r) is at the smallest r-values under these conditions. In other words, the RNA component is contributing more to the total scattering in 0% and 10% D_2O, and this is reflected as a shift in the peak in P(r) to smaller r values. In 85% and 100% D_2O, the RNA component is contributing to the total scattering, but the scattering from the protein component is much stronger. Thus, the peak in P(r) falls in between the 65% D_2O case and the 0% and 10% D_2O cases.

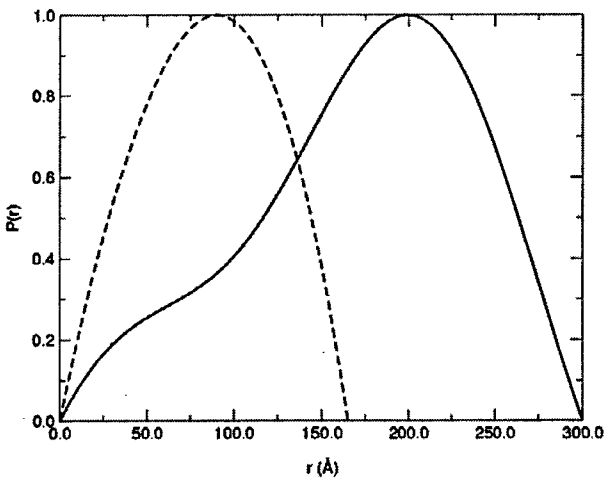

Figure 3. Distance Distribution Functions for the Components of MS2. Distance distribution functions, P(r) vs. r, for the protein (—) and RNA (– – –) components of the MS2 particles from Experiment #1.

The scattered intensities from the RNA and protein components were separated from each other using eq. 2. The resultant P(r) functions are shown in Figure 3 for Experiment #1. This process was also attempted using the data from Experiment #2. However, the results were much noisier, probably owing to the poorer quality of the data relative to the data obtained in Experiment #1. The results for Experiment #1 show that, while D_{max} for the protein shell remains at 300Å, D_{max} for the RNA core was 165Å. Thus, the RNA component appears to be confined mostly within a radius of ~83Å. The peak of the RNA P(r) distribution is also around this value. These results agree very well with the R1 values from the core-shell model fits for the samples measured in 10% D_2O for Experiment #1 and Experiment #2 (see Table 1). Recall, that the 10% D_2O solvent condition is where the RNA scattering is the strongest relative to that of the protein.

4. DISCUSSION

The MS2 bacteriophage is a model organism for a number of important areas of research including viral replication, infection, and assembly (Stockley *et al.*, 1994). Recently, noninfectious, genetically modified forms of the MS2 phage that contain varying amounts of RNA (compared to the wild-type phage) have been developed for use as biological standards (Pasloske et al., 1998; Stockley and Mastico, 2000). These commercially available recombinant particles, Armored RNAs, are used as reference material in research assays for the HIV; Ebola; Borna; Hepatitis A, C, and G; Dengue; Enterovirus; West Nile; and Norwalk viruses, among others (Ambion, personal communication, 2003).

The use of these particles as biological standards in public health screening of humans and livestock has been hampered by the lack of rapid quantitative methods to analyze the physical properties of this family of particles, not found in nature, which cannot be scientifically characterized by traditional methods. These MS2-like biomarkers, because of their small size and the necessity that they be noninfectious, cannot be rapidly or reliably counted. As a result, this new generation of biological reference material cannot be cheaply characterized for general use in public health laboratories. This is solely due to the fact that their physical properties in solution cannot be quantified or confirmed. Thus, there is a need for instrumentation that can count biological particles about which nothing is known and that also can provide structural information about their properties in solution.

The creation of these new forms of MS2 has made it increasingly important to understand the relationship of the indigenous MS2 RNA to its protein shell and to measure the Mw of the wild-type RNA molecule *in vivo* under biological conditions. Currently, this can only be accomplished by combining small-angle neutron scattering (SANS) with a novel virus counting instrument, the IVDS, since there is no other method to rapidly count small (<100 nm) biological particles with unknown properties in solution in the absence of information about the particle molecular weight (Mw) or infectivity. This initial study, which is the first of a larger set of completed experiments using recombinant biomarkers, serves as a model for the use of SANS and IVDS as a virus identification and characterization tool.

Our results show that (a) the MS2 RNA is tightly compacted within the virion, confined to a radius of 83 ± 1Å, (b) the MS2 coat protein shell extends to a radius of 136 ± 2Å and has a thickness of 21 ± 1Å. Significantly, we were able to measure the spatial extent of the protein shell and RNA core separately as well, and to show that the RNA core is compact. For about 40 years, researchers studying MS2 have speculated about the nature of the native MS2 RNA *in vivo* (Strauss and Sinsheimer, 1963). Interest in this issue was initially spurred by a conflict between the ideal Mw of small RNA viruses in general, as deduced from the apparent particle size from electron microscopic studies and the density of the particle in cesium chloride (CsCl), compared to the experimentally determined Mw as measured using sedimentation gradient experiments and the resulting associated Sedvberg constant (S20,w). A virus like MS2, with an approximate particle size of 23 to 28 nm based on electron microscopy, and with a density in CsCl of 1.41 g/cm^3 would be expected to have a molecular weight on the order of 9.8 x 10^6 g/mol. However, based on experimental sedimentation data, the Mw of MS2 was determined to be 3.6 x 10^6 g/mol. This discrepancy between ideal and experimental molecular weight has been seen in a variety of small RNA viruses (Kaesberg, 1959). Initially, this was thought to be due to some unusual features related to the interaction of CsCl with viral RNA. However, subsequent experiments using deuterated water rather than cesium chloride yielded similar results (Overby et al., 1966). The idea that the core of the virus RNA might contain hollow regions with solvent was proposed to explain these discrepancies, but this theory was not proven (Overby et al., 1966). Our results show definitively that indeed the RNA is tightly packed within the MS2 protein shell. The amount of water in the core was calculated from the fitted scattering length density of the core region for the data at each contrast. The average fraction of water in the core region was found experimentally to be 0.81 ± 0.04.

Indirect genetic and biochemical results hint that a variety of mechanisms may act in concert to fold and compact the MS2 genomic RNA. The complete genome of the MS2 RNA has been determined (Fiers et al., 1976). Sequence analysis of the nucleotide sequence predicts that the MS2 RNA spontaneously folds into a number of important secondary structures (Fiers et al., 1976). Perhaps the most important of these secondary motifs is a stem-loop structure found in the MS2 operator site. The MS2 coat protein has been shown to specifically bind the stem-loop site and to selectively regulate the virus assembly and transcription of the MS2 transcriptional operon (Stockley et al., 1994). Co-crystallographic experiments involving the empty MS2 capsid and a RNA fragment containing the stem-loop region of the MS2 operator site have shown that the MS2 coat protein and MS2 genomic RNA interact at 20 different sites along the MS2 RNA backbone.

The A protein also has been shown to directly bind the MS2 genomic RNA at its 5 ft and 3 ft end. A number of lines of evidence point to the tight association between the A protein and the genomic RNA. First, the A protein and genomic RNA are selectively co-precipitated. Second, RNAse protection and competition experiments followed by sequence analysis of the protected RNA indicate that the MS2 RNA is tightly bound by the A protein at its 5 ft and 3 ft end (Shiba and Suzuki, 1981). This tight association between the genomic RNA and the A protein has been shown to be important for RNA packing *in vitro*. Mutant MS2 bacteriophage, which lack the A protein, contain loosely packed RNA that protrudes from the MS2 protein shell and then becomes degraded by nucleases in the media (Argetsinger and Gussin, 1966; Heisenberg, 1966). In addition, the small polyanion, spermidine, is selectively

sequestered in the capsid core (Leipold, 1977; Jacobson et al., 1985). *In vitro* experiments using isolated MS2 RNA/A protein complexes have revealed that increased levels of spermidine lead to increased compaction of the RNA/A protein complex as measured by its migration in sedimentation gradients, as well as by its increased infectivity of the complex *in vitro* (Leipold, 1977).

Taken together, this body of *in vitro* experiments suggests strongly that the MS2 RNA is tightly compacted. Our work is the first study to directly measure the spatial distribution of the MS2 genomic RNA under indigenous conditions and to confirm that it is indeed compact *in vivo*

LITERATURE CITED

1. Argetsinger, J.E., and Gussin, G.N. (1966) Intact Ribonucleic Acid from Defective Particles of Bacteriophage R17. *Journal of Molecular Biology* **21**: 421-434.

2. Atkins, J.F., Steitz, J.A., Anderson, C.W., and Model, P. (1979) Binding of Mammalian Ribosomes to MS2 Phage RNA Reveals an Overlapping Gene Encoding a Lysis Function. *Cell* **18**: 247-256.

3. Caspar, D.L., and Klug, A. (1962) Physical Properties in the Construction of Regular Viruses. *Cold Spring Harbor Symposia Quantitative Biology* **27**: 1-24.

4. Cuillel, M., Tripier, F., Braunwald, J., and Jacrot, B. (1979) A Low Resolution Structure of Frog Virus 3. *Virology* **99**: 277.

5. Cusack, S., Ruigrok, R.W., Krygsman, P.C., and Mellema, J.E. (1985) Structure and Composition of Influenza Virus. A Small Angle Neutron Scattering Study. *Journal of Molecular Biology* **186**: 565-582.

6. Davis, J.E., and Sinsheimer, R.L. (1963) The Replication of Bacteriophage MS2. *Journal of Molecular Biology* **6**: 203-207.

7. Eisenberg, D. (1979) *Physical Chemistry with Applications to the Life Sciences*; Benjamin/Cummings Publishers.

8. Fiers, W., Contreras, R., Duerinck, F., Haegman, G., Iserentant, D., Merregeart, J. *et al.* (1976) Complete Nucleotide Sequence of Bacteriophage MS2 RNA: Primary and Secondary Structure of the Replicase Gene. *Nature* **260**: 500-507.

9. Freeman, R. and Leonard, K.R. (1980) Comparative Mass Measurement of Biological Macromolecules by Scanning Transmission Electron Microscopy. *Journal of Microscopy* **122**.

10. Glinka, C.J., Barker, J.B., Hammouda, B., Krueger, S., Moyer, J.J., and Orts, W.J. (1998) The 30m Small-Angle Neutron Scattering Instruments at the National Institute of Standards and Technology. *Journal of Applied Crystrallography* **31**: 430-445.

11. Golmohammadi, R., Valegard, K., Fridborg, K., and Liljas, L. (1993) Refined Structure of Bacteriophage MS2 at 2.8 Å Resolution. *Journal of Molecular Biology* **234**: 620-639.

12. Heisenberg, M. (1996) Formation of Defective Bacteriophage Particles by Fr. Amber Mutants. *Journal of Molecular Biology* **17**: 136-144.

13. Jacrot, B., and Zaccai, G. (1981) Determination of Molecular Weight by Neutron Scattering. *Biopolymers* **20**: 2413.

14. Mazzone, H. (1998) *CRC Handbook of Viruses: Mass-Molecular Weight Values and Related Properties.* CRC Press LLC.

15. Pasloske, B.L., Walkerpeach, C.R., Obermoeller, R.D., Winkler, M., and DuBois, D.B. (1998) Armored RNA Technology for Production of Ribonuclease-Resistant Viral RNA Controls and Standards. *Journal of Clinical Microbiology* **36**: 3590-3594.

16. Peabody, D.S., and Al-Bitar, L. (2001) Isolation of Viral Coat Protein Mutants with Altered Assembly and Aggregation Properties. *Nucleic Acids Research* **29**: e113.

17. Sambrook, J., and Russell, D.W. (2001) Molecular Cloning a Laboratory Manual. *Cold Spring Harbor Press: Cold Spring Harbor, New York,* 2.38-32.39, 32.44, 32.48-32.51.

18. Semenyuk, A.V., and Svergun, D.I. (1991) GNOM- a Program Package for Small-Angle Scattering Data Processing. *Journal of Applied Crystrallography* **24**: 537-545.

19 Shiba, T., and Suzuki, Y. (1981) Localization of a Protein in the RNA-A Complex of RNA Phage MS2. *Biochimica et Biophysica Acta* **654**: 249-255.

20. Stockley, P.G., Stonehouse, N.J,. and Valegard, K. (1994) Molecular Mechanism of RNA Phage Morphogenesis. *International Journal of Biochemistry.* **26**: 1249-1260.

21. Stockley, P.G., and Mastico, R.A. (2000) Use of Fusions to Viral Coat Proteins as Antigenic Carriers for Vaccine Development. Methods in Enzymology 326: 551-569.

22. Strauss, J.H., and Sinsheimer, R.L. (1963) Purification and Properties of Bacteriophage MS2 and of its Ribonucleic Acid. *Journal of Molecular Biology* **7**: 43-54.

23. Struhrmann, H.B., and Miller, A. (1978) Small-Angle Neuton Scattering of Biological Structures. *Journal of Applied Crystrallography* **11**: 325-345.

24. Svergun, D.I., and Koch, M.H.J. (2002) Advances in Structure Analysis Using Small-Angle Scattering in Solution. *Current Opinion in Structural Biology* **12**: 654-660.

25. Torbet, J. (1979) Neutron Scattering Study of the Solution Structure of Bacteriophages PF1 and Fd. *FEBS Letter* **108**: 61.

26. Valegard, K., Liljas, L., Fridborg, K., and Unge, T. (1990) Three Dimensional Structure of the Bacterial Virus MS2. *Nature* **345**: 36-41.

27. Valegard, K., Liljas, L., Fridborg, K., and Unge, T. (1991) Structure Determination of the Bacteriophage MS2. *Acta Crystallographica* **B47**: 949-960.

28. Valegard, K., Unge, T., Montelius, I., Strandberg, B., and Fiers, W. (1986) Purification, Crystallization and Preliminary X-ray Data of the Bacteriophage MS2. *Journal of Molecular Biology* **190**: 587-591.

29. Wick, C.H., and McCubbin, P.E. (1999a) Purification of MS2 Bacteriophage from Complex Growth Media and Resulting Analysis by the Integrated Virus Detection System (IVDS). *Toxicological Methods* **9**: 253-263.

30. Wick, C.H., and McCubbin, P.E. (1999b) Passage of MS2 Bacteriophage through Various Molecular Weight Filters. *Toxicological Methods* **9**: 265-273.

31. Wick, C.H., McCubbin, P.E. (1999c) Characterization of Purified MS2 Bacteriophage by the Physical Counting Methodology used in the Integrated Virus Detection System (IVDS). *Toxicological Methods* 1999c **9**: 245-252.